CORRELACIONES EN FINANZAS

Juan Carlos García Céspedes

E-mail jcgarcia@grupobbva.com

SEMINARIO DE MATEMATICAS FINANCIERAS

Instituto MEFF - UAM

Enero 2.001

OBJETIVOS

- Discutir el concepto de correlación
- Exponer algunas limitaciones.
- Presentar algunas herramientas y conceptos utiles.
- Todo ello en relación con las finanzas (riesgo de mercado y riesgo de crédito).

CORRELACION

- Concepto absolutamente extendido y a la vez mal entendido

- En general por correlación se entiende cualquier tipo de dependencia/relación entre variables aleatorias.

- Sin embargo, la correlación es tan sólo una medida entre muchas otras de la relación que existe entre variables aleatorias.

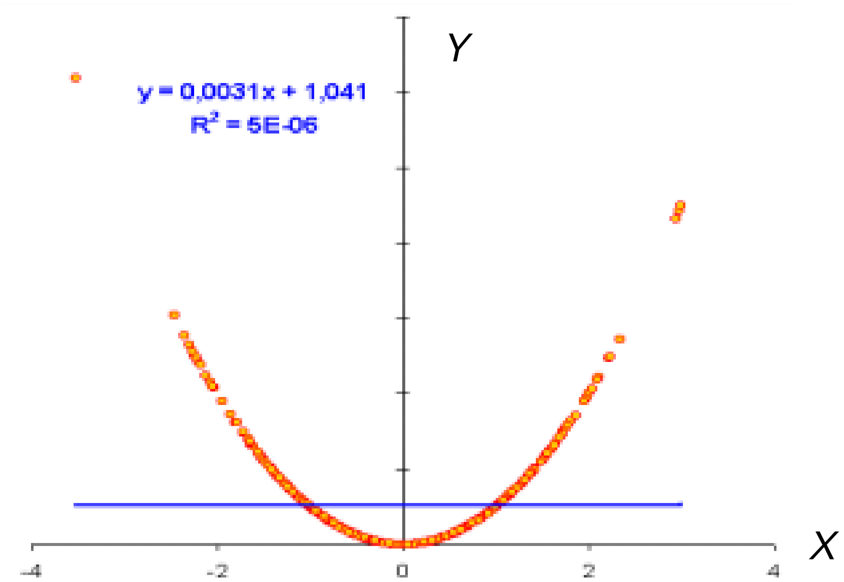
- Es la medida base en el universo de las distribuciones normales (gaussianas), mas genéricamente en el universo de las distribuciones esféricas y elípticas.

- Se ha extendido por extrapolación la utilidad de la correlación lineal al resto de distribuciones

ALGUNOS PROBLEMAS CON LA CORRELACIÓN:

✓ Para que exista correlación, las varianza de X e Y deben existir, esto es, las variables aleatorias con distribuciones de colas gruesas no tienen correlación

✓ Independencia implica correlación cero, pero no al revés. Ejemplo $Y=X^2$, donde X es normal (0,1).



ALGUNOS PROBLEMAS MÁS CON LA CORRELACIÓN:

¿¿¿Existe alguna contradicción????

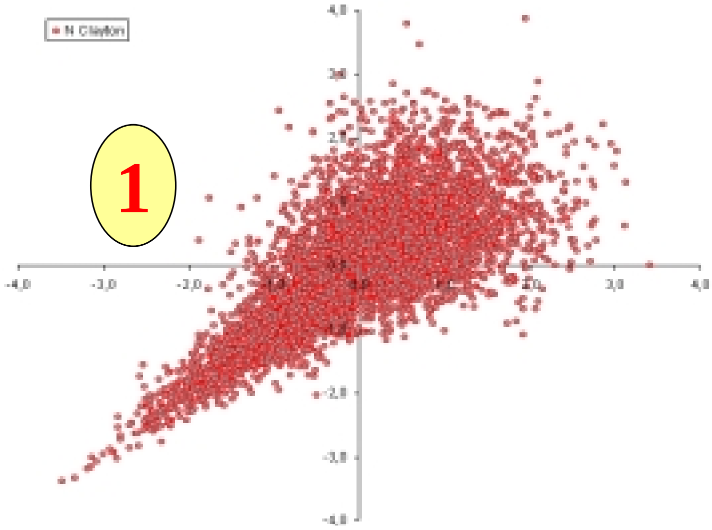
- ✓ Solo en el entorno de las normales multivariantes podemos interpretar correlación cero igual a independencia.
- ✓ Sin embargo, dos normales correlacionadas cero no implica necesariamente independencia

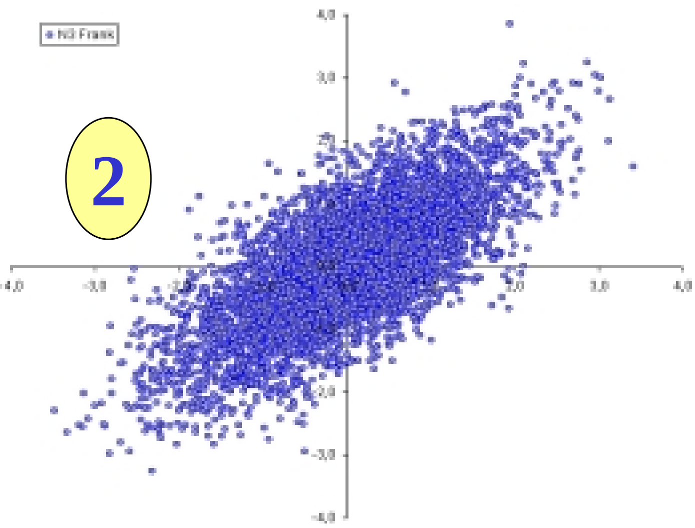
NO, porque…

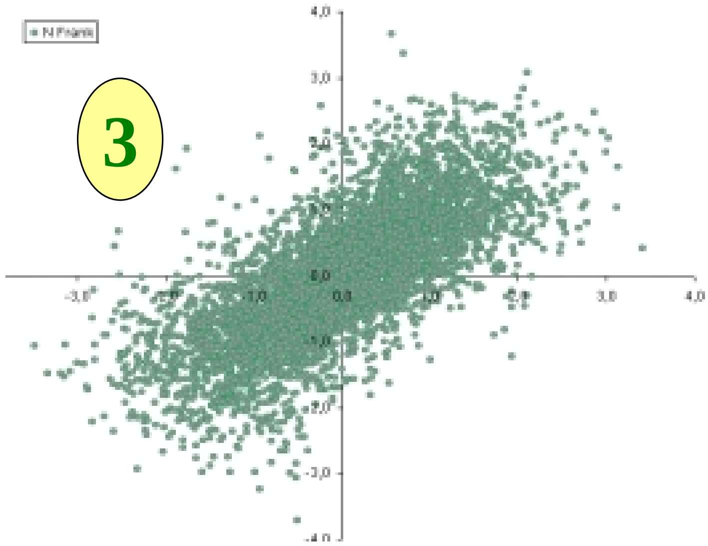
- ✓Que las distribuciones marginales sean Normales no implica que la distribución conjunta sea una normal multivariante,
- ✓Existen infinitas distribuciones bivariantes con una correlación dada y con marginales gaussianas. Veamos un ejemplo...

CUESTIÓN

Uno de estos conjuntos de puntos sigue una distribución normal bivariante. ¿Cuál?

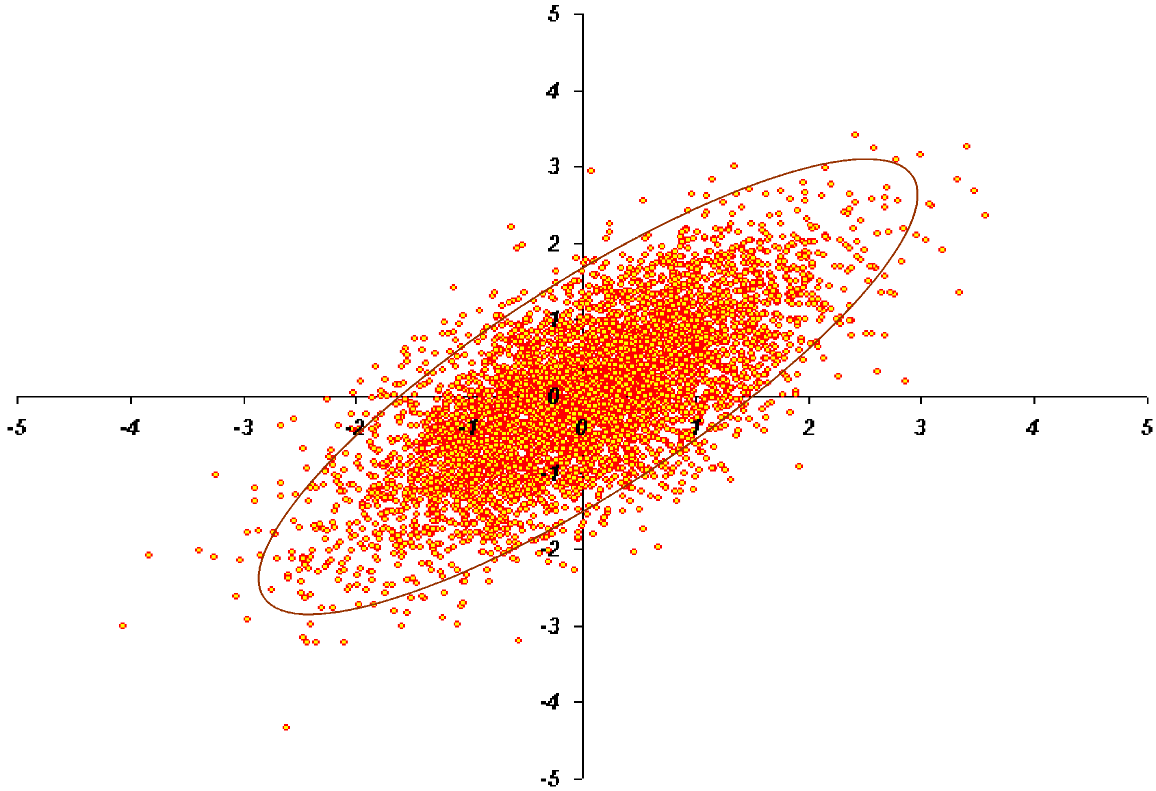




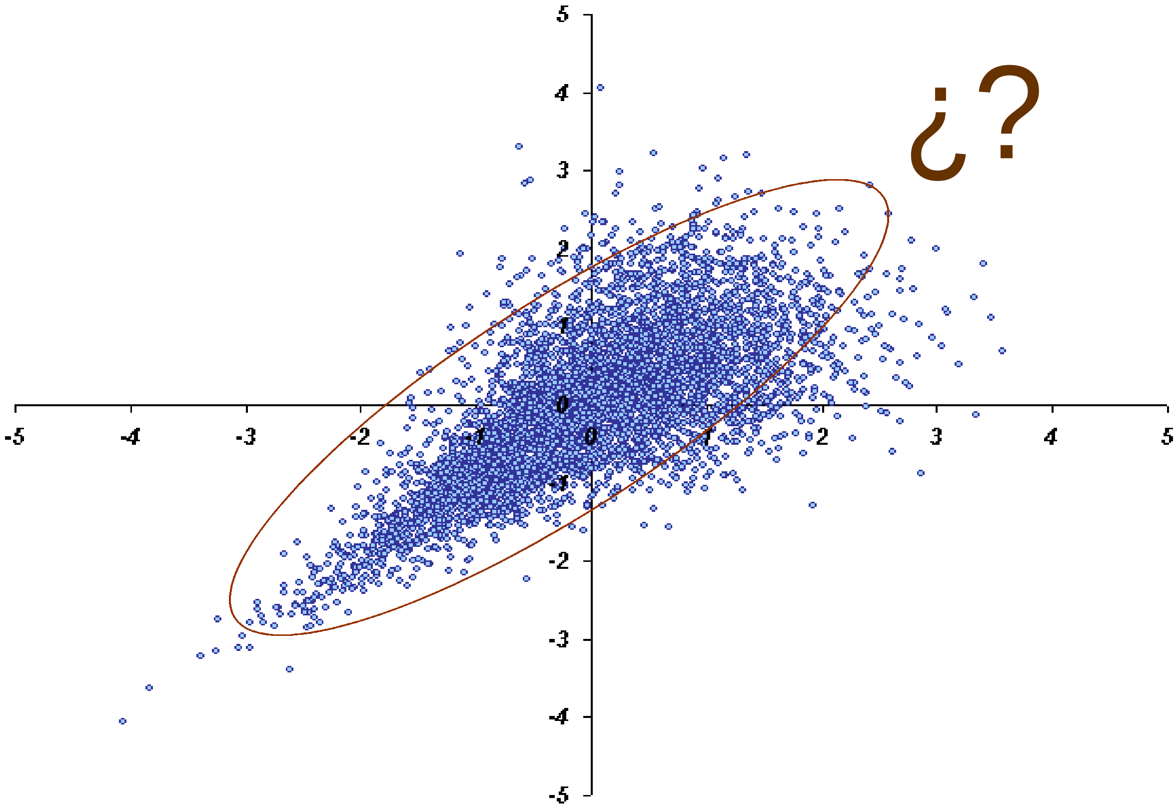


Vamos a quedarnos con las distribuciones 1 y 2

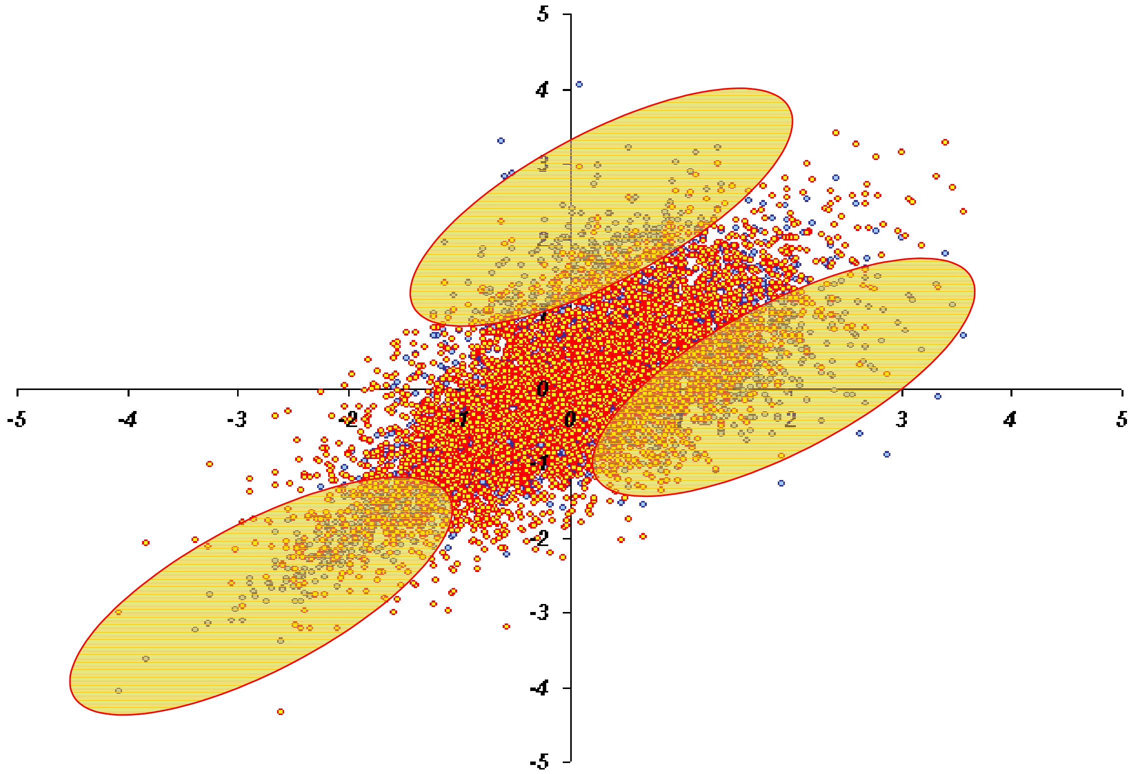
Estos puntos se han generado a partir de dos distribuciones normales correlacionadas (ρ = 0,71)



Estos otros puntos... ¡También se han generado a partir de dos distribuciones normales correlacionadas (ρ = 0,71)!

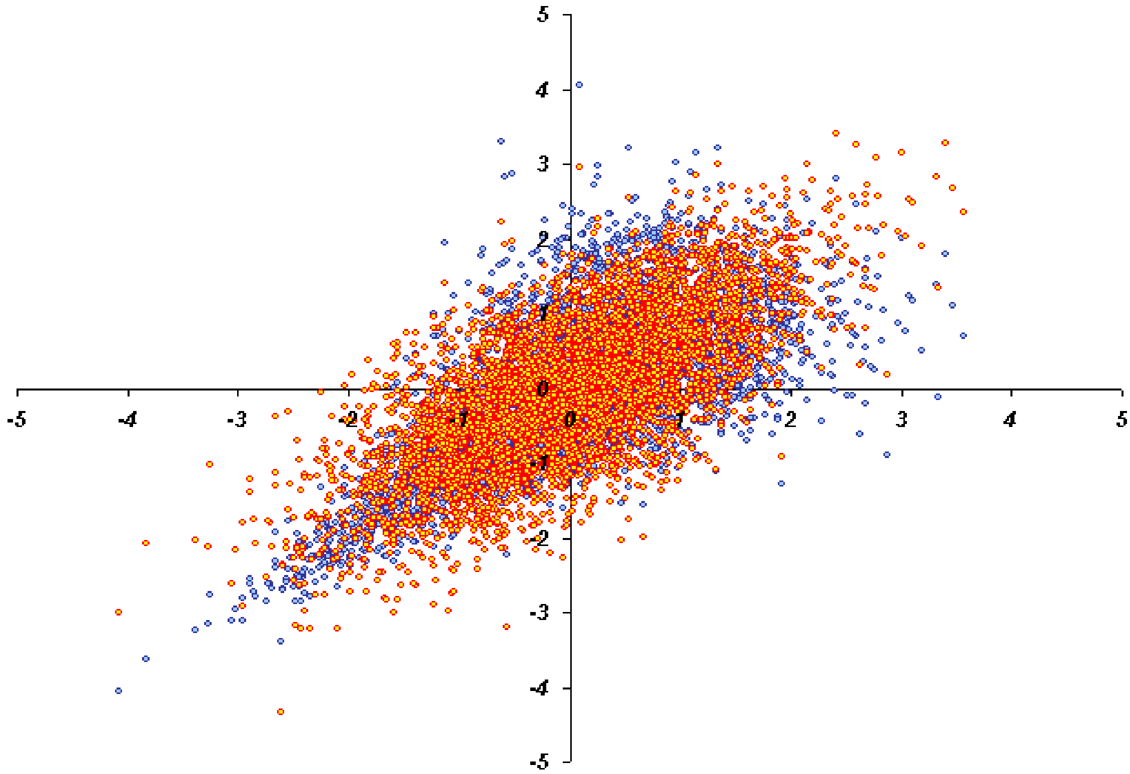


Las diferencias entre los dos conjuntos de puntos son evidentes...



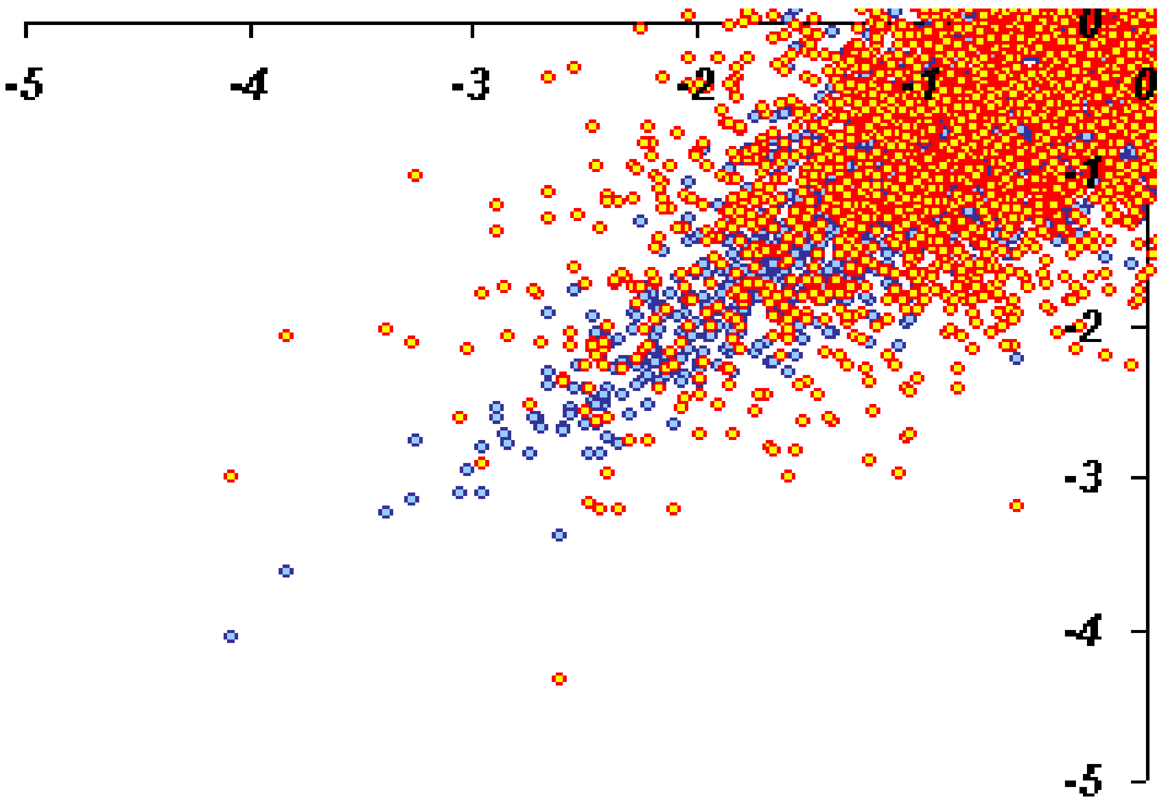
...no siguen la misma distribución y sin embargo sus distribuciones marginales son normales con la misma correlación.

Las diferencias entre los dos conjuntos de puntos son evidentes...

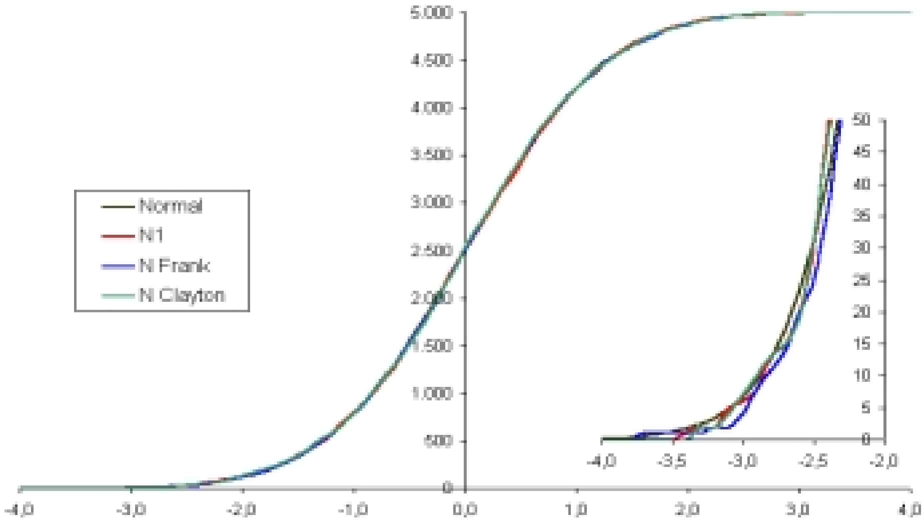


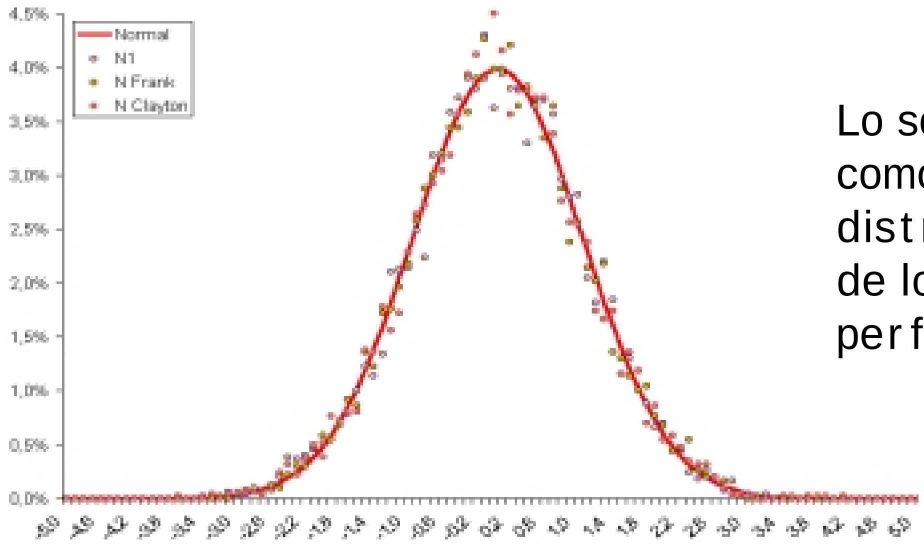
...no siguen la misma distribución y sin embargo sus distribuciones marginales son normales con la misma correlación.

¿qué conclusiones podemos sacar en términos de riesgo de mercado?



... Aunque los factores de riesgo sean individualmente normales, el riesgo de la cartera puede ser mayor de lo que creemos.





Lo sorprendente es que, como vemos, las distribuciones marginales de los ejemplos son perfectamente normales

LAS CÓPULAS (en estadística)

Una cópula es una distribución multivariante cuyas distribuciones marginales son uniformes (0,1).

$$C\left(u_1, u_2, ..., u_n\right)$$

Sean $x_1, x_2, ..., x_n$ variables aleatorias con distribuciones acumuladas marginales $F_1(x_1), F_2(x_2), ..., F_n(x_n)$. Entonces:

$$F\left(x_1, x_2, ..., x_n\right) = C\left(F_1(x_1), F_2(x_2), ..., F_n(x_n)\right)$$

define una distribución conjunta acumulada cuyas marginales son precisamente $F_1(\cdot), F_2(\cdot), ..., F_n(\cdot)$.

LAS CÓPULAS (en estadística)

Es posible mapear las variables aleatorias en variables uniformes a través de su distribución acumulada. La estructura de dependencia vendrá determinada por relaciones entre distribuciones uniformes.

Si se le da la vuelta al argumento, una cópula es un instrumento magnífico para la simulación de variables aleatorias con distribuciones marginales dadas, solo habremos de simular variables uniformes con estructuras de correlaciones determinadas.

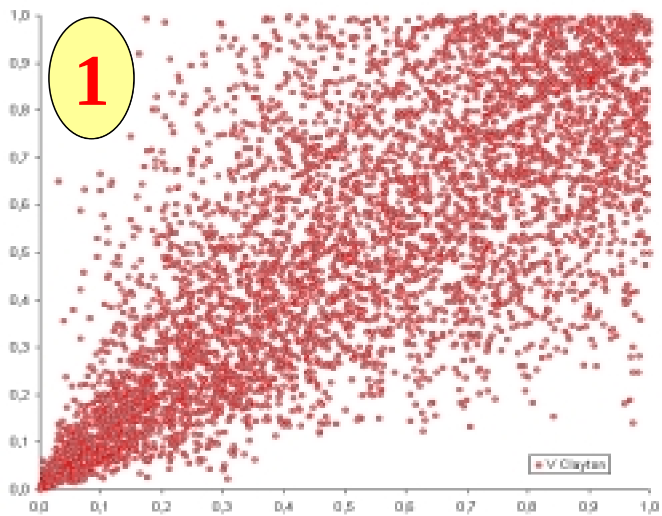
Teorema de Sklar

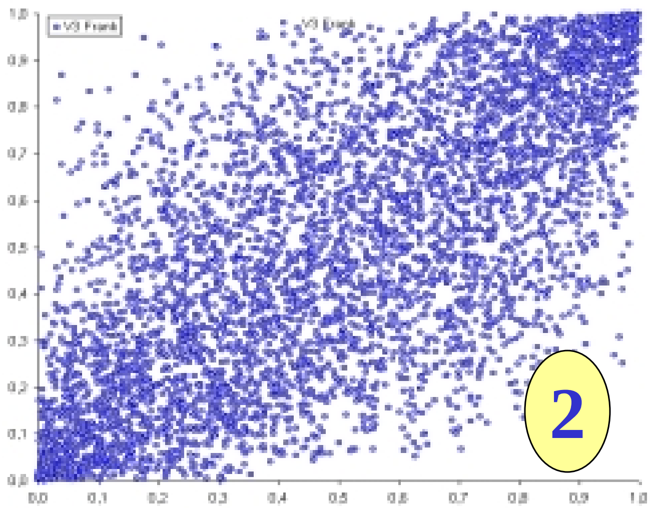
Sea $F(x_1,x_2,\ldots,x_n)$ una función de distribución acumulada, entonces existe una n-cópula tal que,

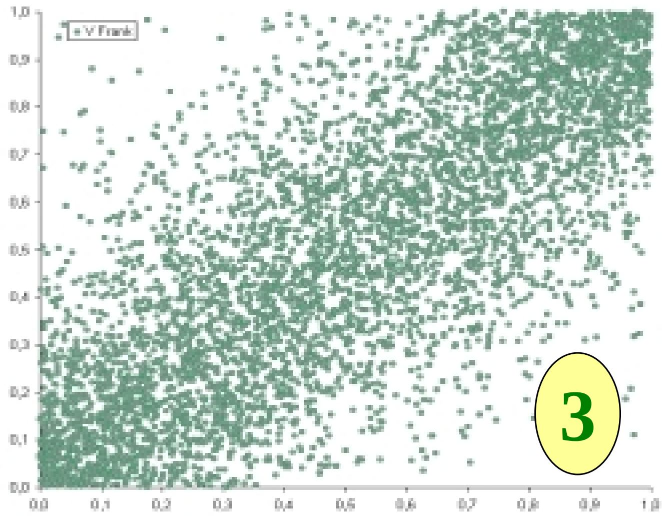
$$F\left(x_1, x_2, \ldots,\ x_n\right) = C\left(F_1(x_1), F_2(x_2), \ldots,\ F_n(x_n)\right)$$

Las copulas permiten separar la estructura de dependencia multivariante de la estructura de distribuciones marginales univariantes.

Veamos que sucede si mapeamos las distribuciones anteriores a uniformes...



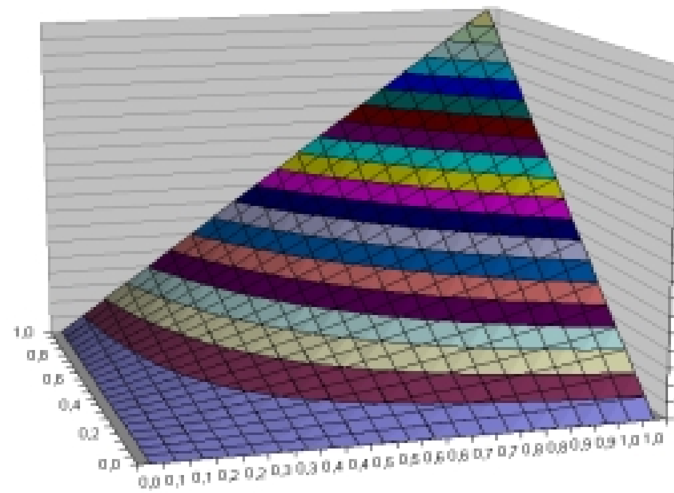


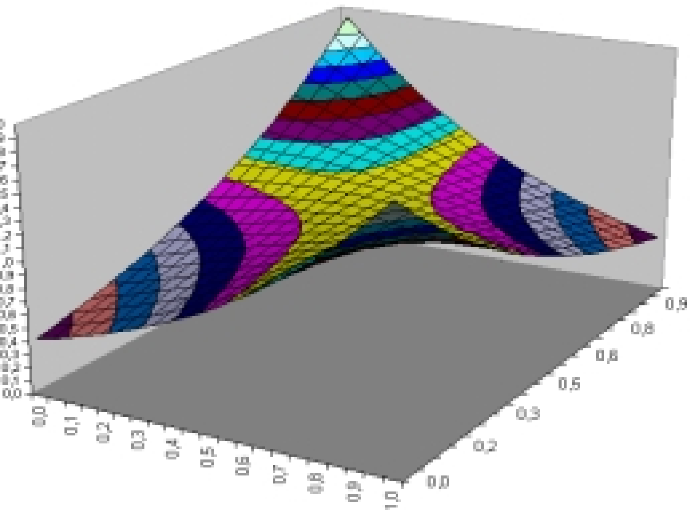


El gráfico 3 se ha generado a partir de la llamada **cópula de Frank**.
A continuación podemos ver su distribución acumulada y su densidad.

$$C(u,v) = \frac{1}{\alpha} \cdot \ln\left[1 + \frac{\left(e^{\alpha \cdot u} - 1\right)\cdot\left(e^{\alpha \cdot v} - 1\right)}{e^{\alpha} - 1}\right] \qquad -\infty < \alpha < \infty$$

Cuanto mayor es el parámetro α, en valor absoluto, mayor es la relación entre las dos variables.

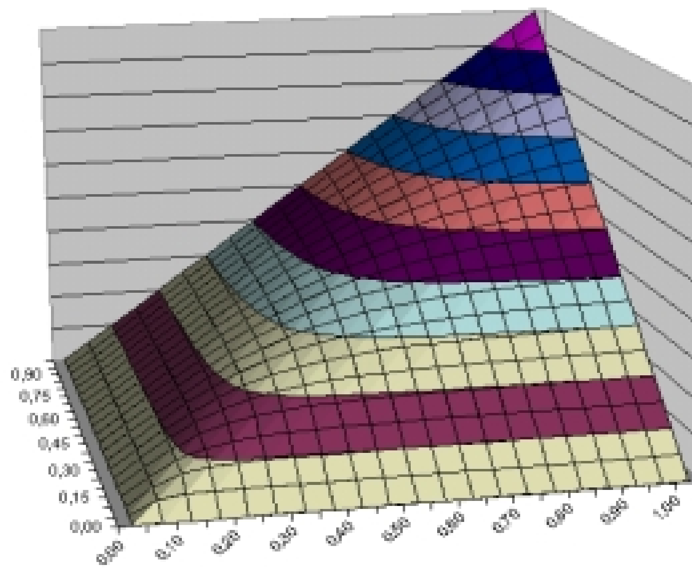


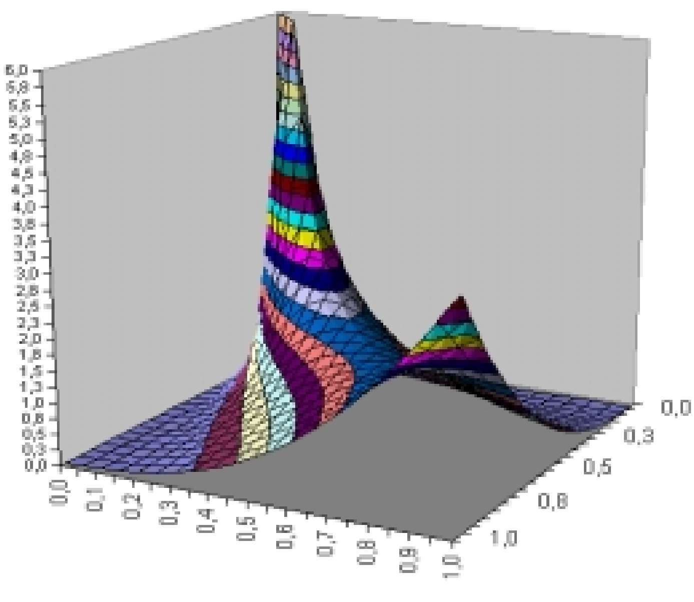


El gráfico 1 se ha generado a partir de la llamada **cópula de Clayton**.
A continuación podemos ver su distribución acumulada y su densidad.

$$C(u,v) = \left(u^{-\alpha} + v^{-\alpha} - 1\right)^{-\frac{1}{\alpha}} \qquad \alpha > 1$$

Nuevamente, cuanto mayor es el parámetro α, mayor es la relación entre las dos variables.





¿Qué está pasando? ¿Qué conclusión podemos sacar?

El concepto de correlación “hace aguas”, sólo tiene sentido en determinados entornos, en otros hay que interpretarlo con sumo cuidado.

En resumen: **La correlación no caracteriza la estructura de relaciones entre las variables aleatorias, estas estructuras en general son muy ricas y difícilmente se pueden capturar con un solo parámetro.**

Existen otras medidas de correlación, o mejor dicho, de relación entre variables aleatorias, mencionemos dos:

- La correlación de Spearman
- La Tau de Kendall

Recordemos el concepto “tradicional” de correlación, que llamaremos **correlación lineal**:

$$\rho_l(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\cdot\sqrt{Var(Y)}}$$

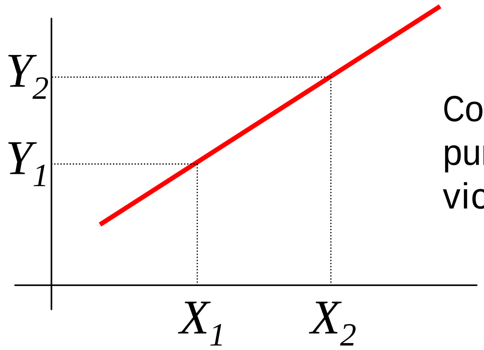
La correlación de spearman se define como:

$$\rho_s(X,Y) = \rho_l(F_X(X),F_Y(Y)) = \frac{Cov(F_X(X),F_Y(Y))}{\sqrt{Var(F_X(X))}\cdot\sqrt{Var(F_Y(Y))}}$$

La correlación de spearman no es más que la correlación lineal pero en el mundo de las distribuciones uniformes, esto es, las variables aleatorias previamente se transforman en uniformes.

La **Tau de Kendall** se define como:

$$\rho_\tau = \underbrace{\mathrm{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right]}_{\text{Puntos concordantes}} - \underbrace{\mathrm{P}\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right]}_{\text{Puntos discordantes}}$$



Correlación positiva indica que hay más puntos concordantes que discordantes y viceversa

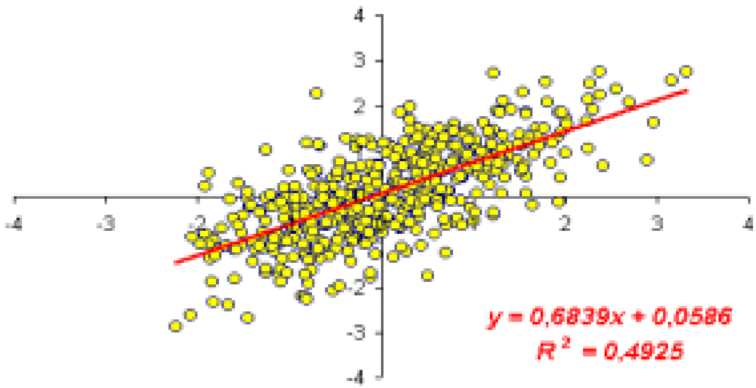
Existe una relación entre la corre-lación lineal y la Tau de Kendall

$$\rho_\tau(X,Y) = \frac{2}{\pi} \cdot \arcsin\left(\rho_l(X,Y)\right)$$

Para estimar empíricamente la Tau de Kendall se recomienda la fórmula siguiente

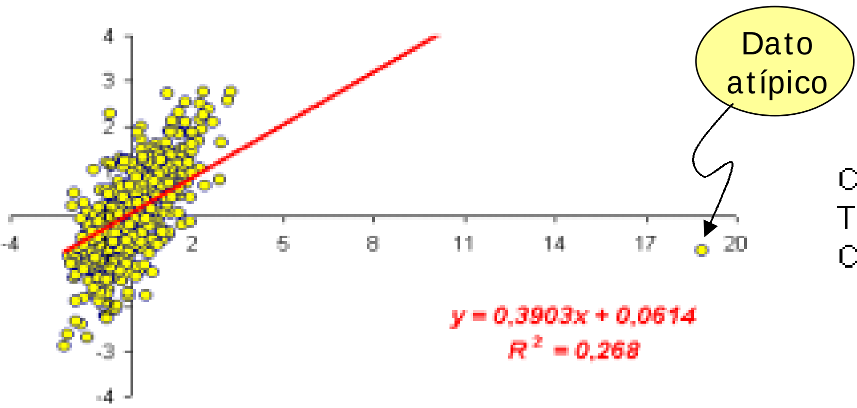
$$\hat{\tau} = \frac{c - d}{\sqrt{c + d + e_X} \cdot \sqrt{c + d + e_Y}}$$

correlación lineal se ve muy afectada por los datos atípicos...



MUESTRA SIN ATIPICOS

Correlación lineal	70,23%
Tau de Kendall	0,49996
Corr. a partir de Tau	70,71%



MUESTRA CON ATIPICOS

Correlación lineal	51,77%
Tau de Kendall	0,49488
Corr. a partir de Tau	70,14%

...sin embargo la Tau de Kendall es mucho mas robusta

Riesgo de crédito, correlaciones y la probabilidad conjunta de incumplimiento

Para dos préstamos existen cuatro posibles escenarios:

Sólo incumple 2 (D_2=1, D_1=0)	**Incumplen ambas** (D_2=1, D_1=1)
Ninguna incumple (D_2=0, D_1=0)	**Sólo incumple 1** (D_2=0, D_1=1)

p_1: Probabilidad de incumplimiento de la compañía 1. $P(D_1=1)$

p_2: Probabilidad de incumplimiento de la compañía 2. $P(D_2=1)$

$p_{1\&2}$: Probabilidad de incumplimiento conjunto. $P(D_1=D_2=1)$

¡ Nótese que por definición $p_{1\&2} \leq Min(p_1,p_2)$!

En el entorno de las variables aleatorias bernouilli, las correlaciones son “algo” diferentes:

$$Datos:$$

$$E(D_1) = p_1 \qquad \sigma_{D_1} = \sqrt{p_1 \cdot (1-p_1)}$$

$$E(D_2) = p_2 \qquad \sigma_{D_2} = \sqrt{p_2 \cdot (1-p_2)}$$

$$corr(D_1, D_2) = \frac{\operatorname{cov}(D_1, D_2)}{\sigma_{D_1} \cdot \sigma_{D_2}}$$

$$\operatorname{cov}(D_1, D_2) = (p_1 - p_{1\&2}) \cdot (1-p_1) \cdot (0-p_2) + (p_2 - p_{1\&2}) \cdot (0-p_1) \cdot (1-p_2) +$$
$$+ p_{1\&2} \cdot (1-p_1) \cdot (1-p_2) + (1-(p_1+p_2-p_{1\&2})) \cdot (0-p_1) \cdot (0-p_2) = \ldots =$$
$$= p_{1\&2} - p_1 \cdot p_2$$

$$corr(D_1, D_2) = \rho_{D_1, D_2} = \frac{p_{1\&2} - p_1 \cdot p_2}{\sqrt{p_1 \cdot (1-p_1)} \cdot \sqrt{p_2 \cdot (1-p_2)}}$$

La correlación vemos que es función de las probabilidades individuales de incumplimiento y de la conjunta.

$$corr(D_1,D_2)=\rho_{D_1,D_2}=\frac{p_{1\&2}-p_1\cdot p_2}{\sqrt{p_1\cdot(1-p_1)}\cdot\sqrt{p_2\cdot(1-p_2)}}$$

Vamos a analizar la función anterior, un primer caso, sencillo, es $p_1=p_2=p$

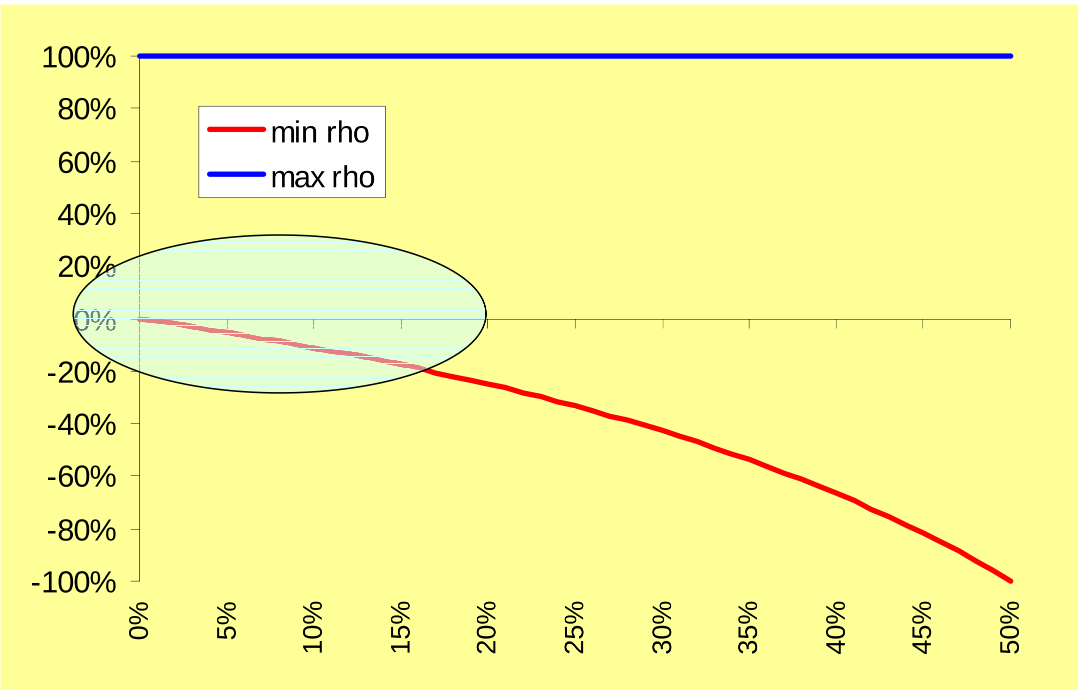
$$corr(D_1,D_2)=\rho_{D_1,D_2}=\frac{p_{1\&2}-p\cdot p}{\sqrt{p\cdot(1-p)}\cdot\sqrt{p\cdot(1-p)}}=\frac{p_{1\&2}-p^2}{p\cdot(1-p)}$$

$$p_{1\&2}=p^2+\rho_{D_1,D_2}\cdot\left(p\cdot(1-p)\right)$$

$$Como\ \ 0\le p_{1\&2}\le p\ \ entonces\ \ \frac{-p}{(1-p)}<\rho_{D_1,D_2}<1$$

¡ La correlación no puede ser -1 salvo que p=0,5 !

Límites a las correlaciones de default entre dos contrapartidas cuando las dos probabilidades de incumplimiento son iguales



Si las probabilidades de default son bajas, que es lo habitual, las correlaciones no pueden muy negativas y la cobertura del riesgo deberá hacerse mediante diversificación.

$$corr(D_1, D_2) = \rho_{D_1,D_2} = \frac{p_{1\&2} - p_1 \cdot p_2}{\sqrt{p_1 \cdot (1-p_1)} \cdot \sqrt{p_2 \cdot (1-p_2)}}$$

El caso general es para $p_1 \neq p_2$: supongamos sin pérdida de generalidad que $p_1 < p_2$

$$p_{1\&2} = p_1 \cdot p_2 + \rho_{D_1,D_2} \cdot \sqrt{p_1 \cdot (1-p_1)} \cdot \sqrt{p_2 \cdot (1-p_2)}$$

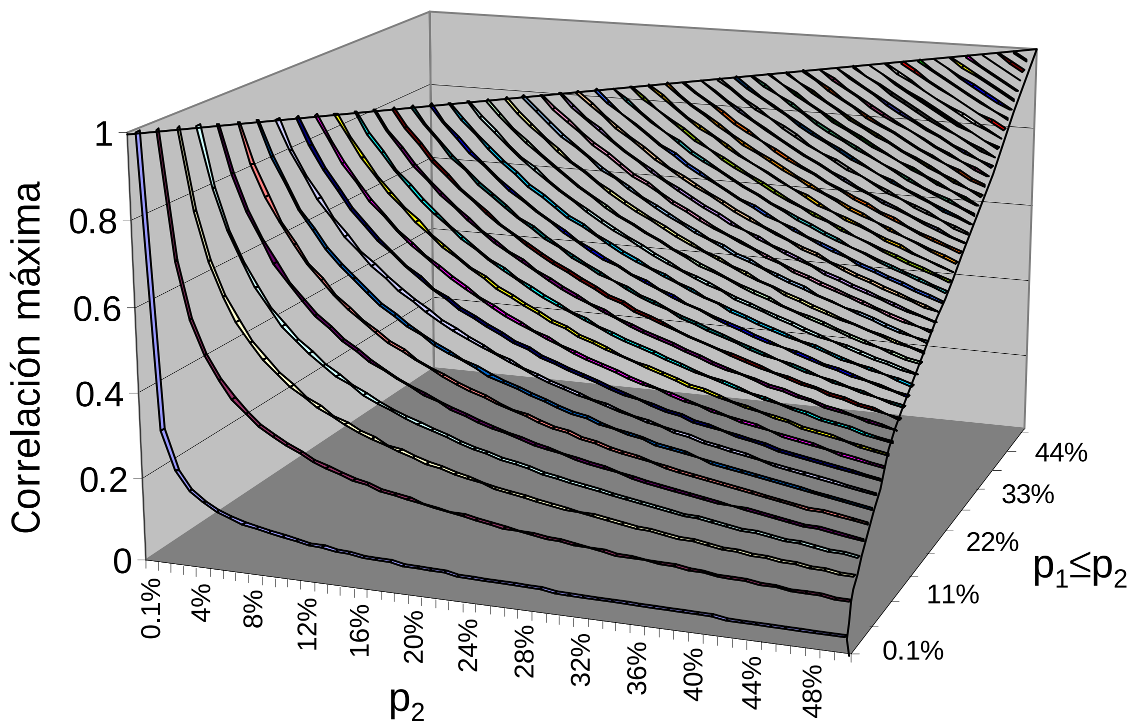
$$Como \;\; 0 \leq p_{1\&2} \leq min(p_1, p_2) = p_1 \;\; entonces$$

$$\frac{-p_1 \cdot p_2}{\sqrt{p_1 \cdot (1-p_1)} \cdot \sqrt{p_2 \cdot (1-p_2)}} < \rho_{D_1,D_2} < \sqrt{\frac{p_1 \cdot (1-p_2)}{p_2 \cdot (1-p_1)}}$$

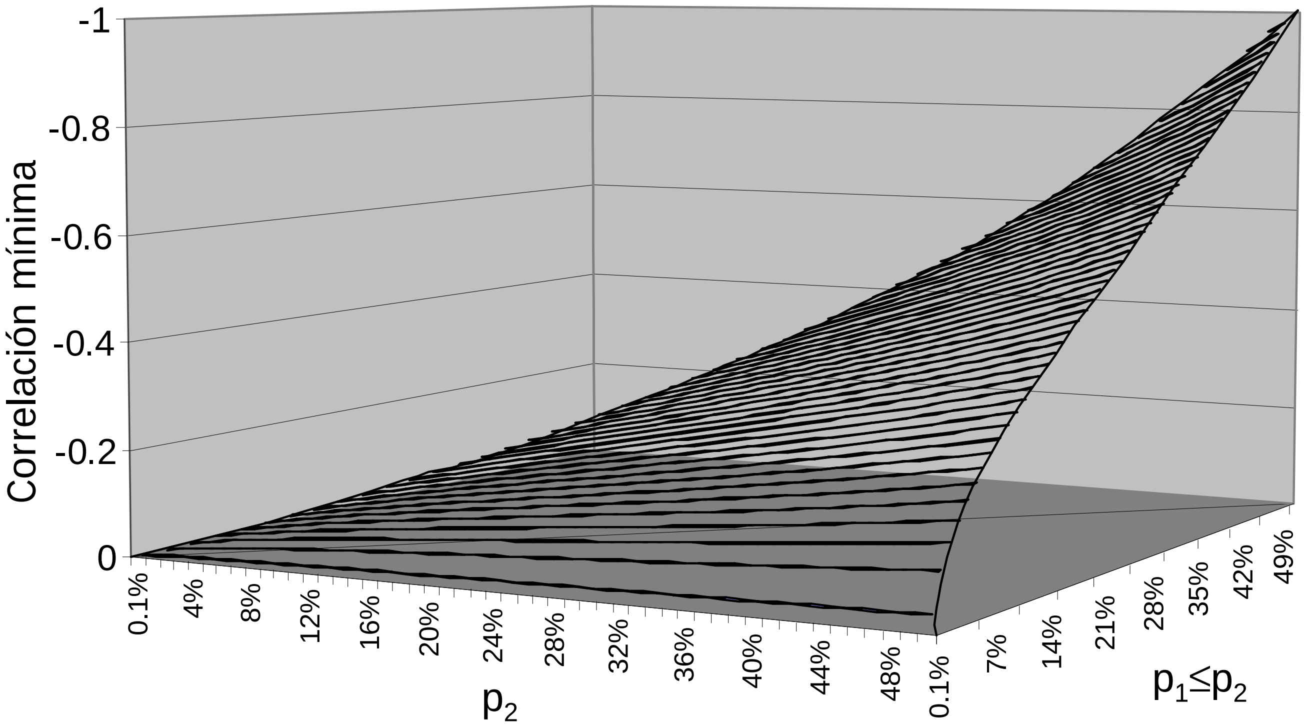
¡ Salvo que $p_1 = p_2$ la correlación no puede ser igual a 1 !

Por tanto, las contrapartidas con diferente calidad crediticia no podrán tener correlación perfecta.

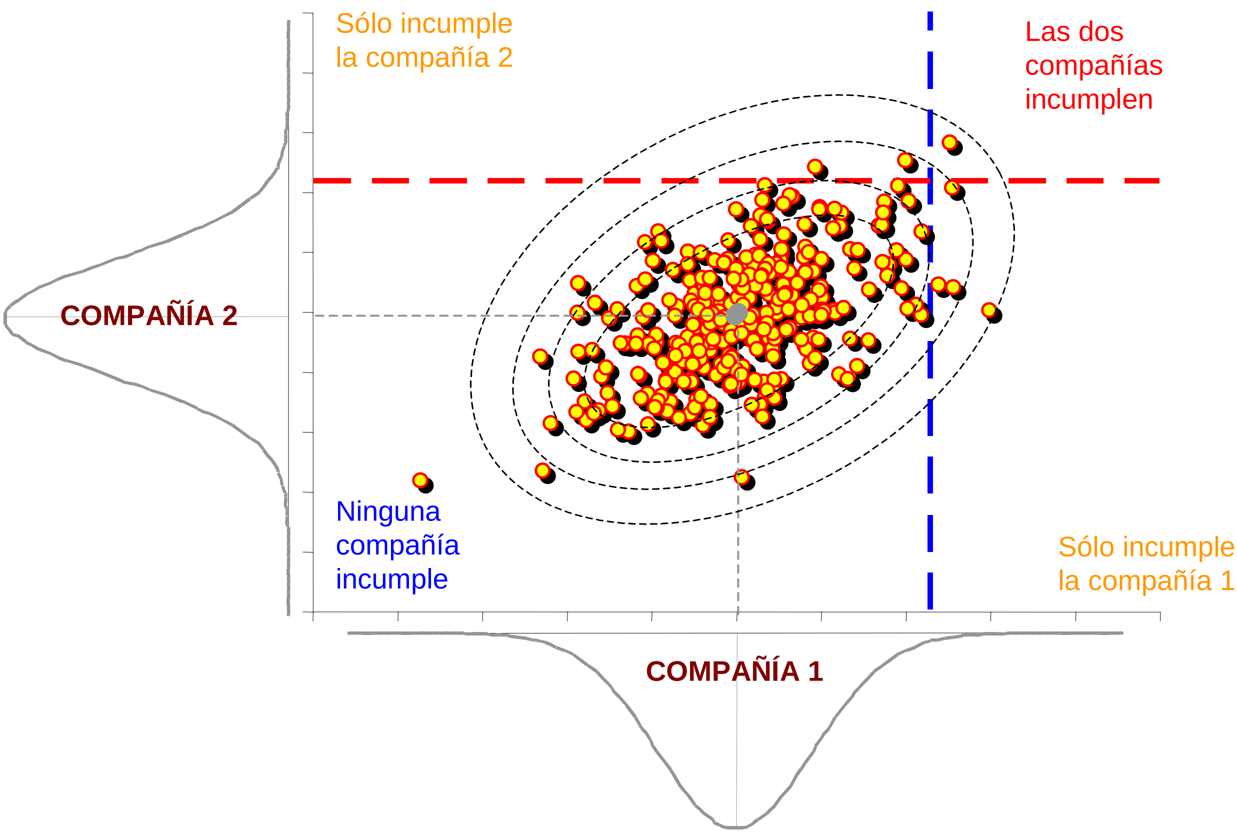
Límite superior de las correlaciones de default en el caso general en el que las probabilidades de incumplimiento no son iguales



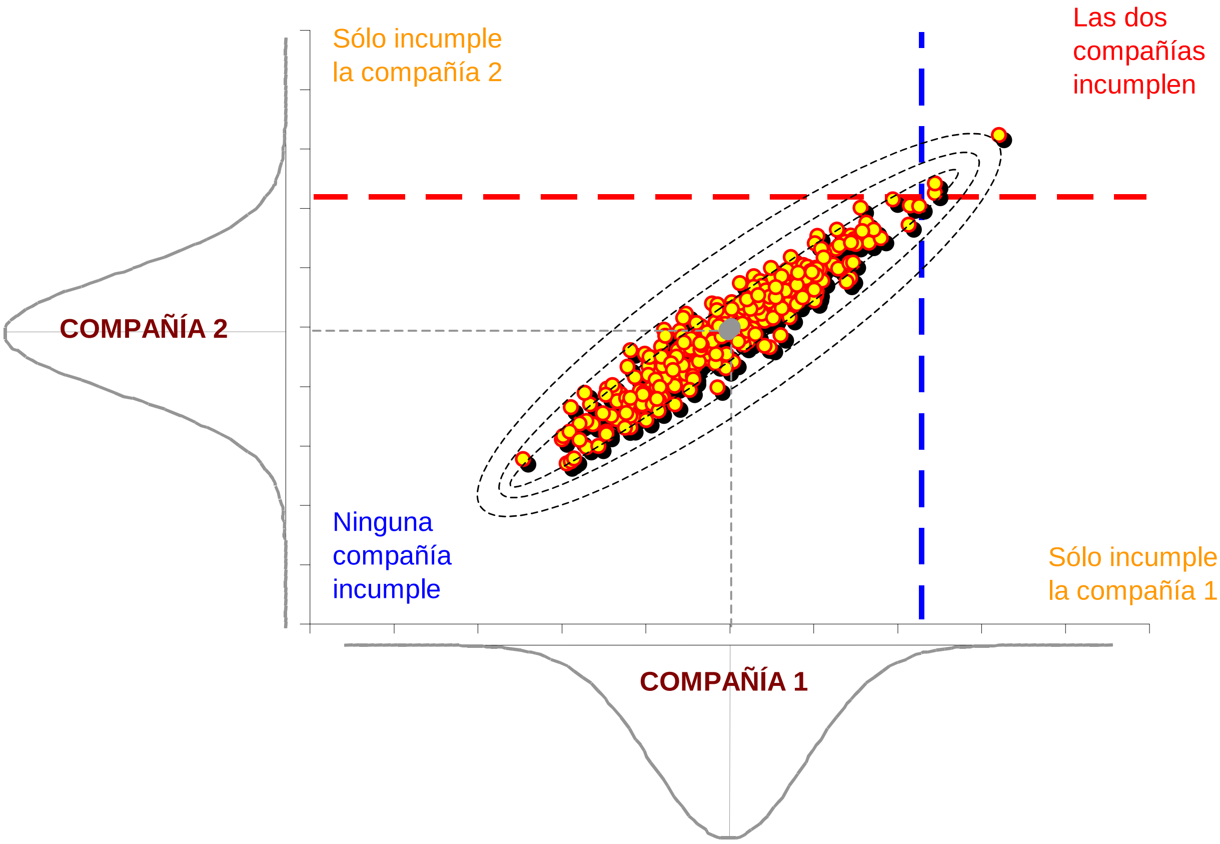
Límite inferior a las correlaciones de default en el caso general en el que las probabilidades de incumplimiento no son iguales



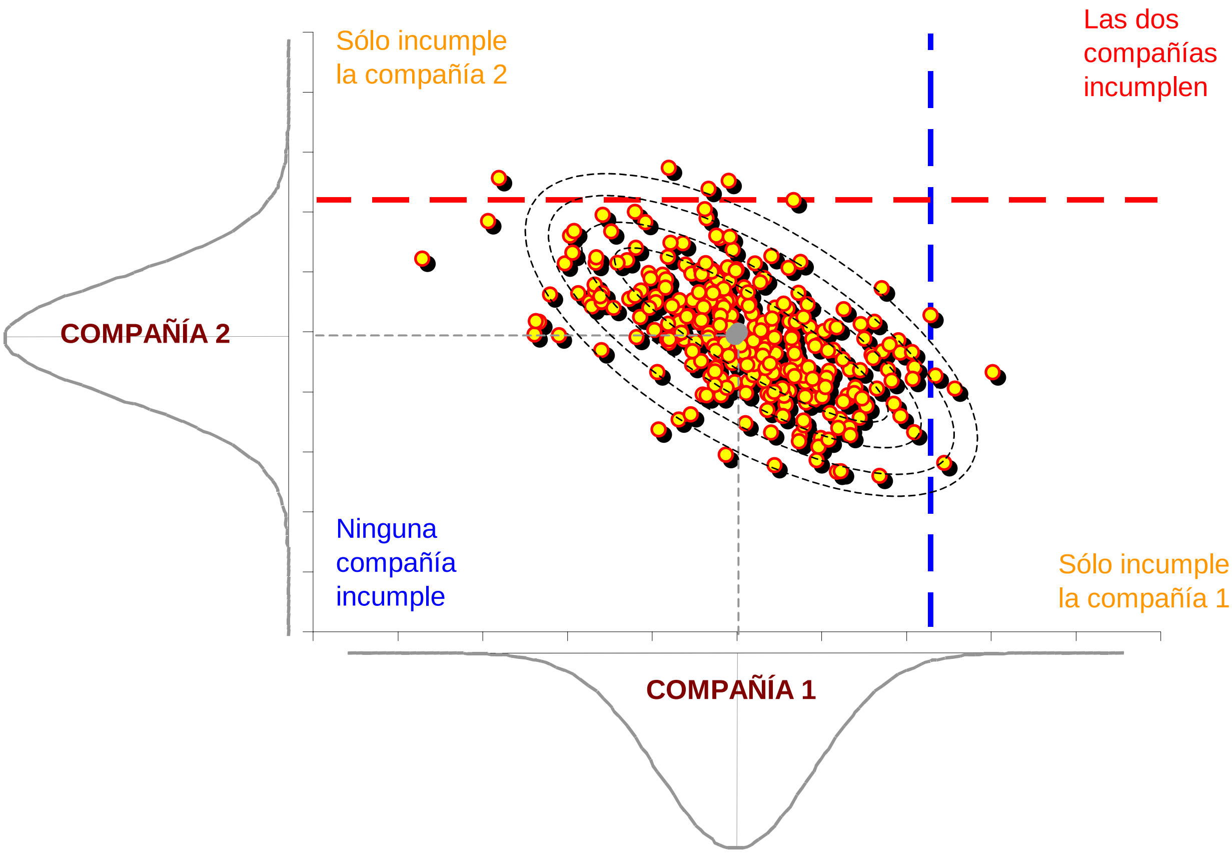
¿Cómo se estima la probabilidad de incumplimiento conjunto? Veamos el modelo de Merton para dos contrapartidas con correlación de activos positiva



El modelo de Merton para dos contrapartidas con correlación de activos cercana a 1

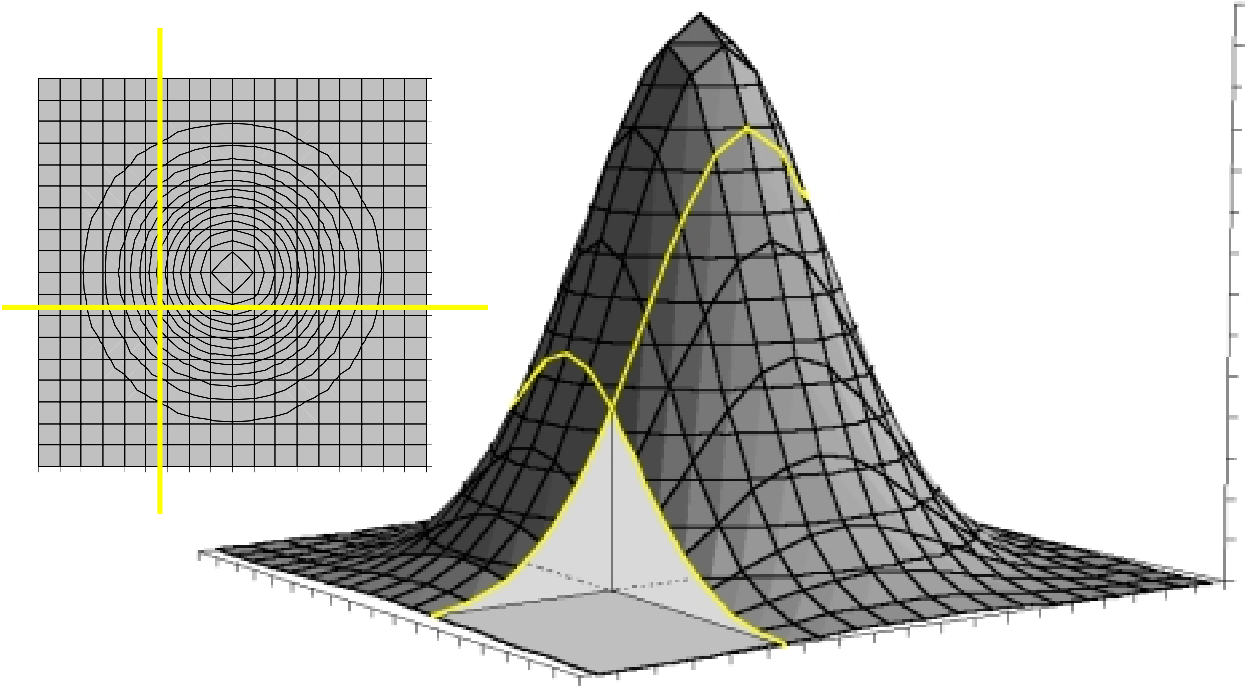


El modelo de Merton para dos contrapartidas con correlación de activos negativa



Cómo se estima la probabilidad de incumplimiento conjunto? Para ello "tan sólo hay que calcular el volumen de campana que cae por encima de la región de incumplimiento conjunto", dicho volumen dependerá de 3 factores:

- Las 2 probabilidades individuales de incumplimiento
- La correlación de activos



Finalmente, vemos que la correlación de default es también una función de los mismos 3 factores:

- Las 2 probabilidades individuales de incumplimiento
- La correlación de activos

$$corr\ (D_1, D_2) = \rho_{D_1, D_2} = \frac{p_{1\&2} - p_1 \cdot p_2}{\sqrt{p_1 \cdot (1 - p_1)} \cdot \sqrt{p_2 \cdot (1 - p_2)}} =$$

$$= \frac{g(p_1, p, \rho_{Activos}) - p_1 \cdot p_2}{\sqrt{p_1 \cdot (1 - p_1)} \cdot \sqrt{p_2 \cdot (1 - p_2)}} = f(p_1, p, \rho_{Activos})$$

El modelo de Merton es un modelo de cópulas semejante a los que antes hemos visto, en este caso se determina la estructura de dependencia entre variables aleatorias bernouilli a través de variables ocultas gaussianas. Es lo que se denomina una "copula gaussiana".

$$C_\Omega(u_1, u_2, \ldots, u_n) = \Phi_\Omega\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)\right)$$

Donde, $\Phi_\Omega(\cdot)$ es una distribución normal estandar multivariante con estructura de correlaciones Ω y $\Phi^{-1}(\cdot)$ son normales univariantes estandar inversas

RESUMEN Y CONCLUSIONES

- Hemos visto que la correlación es un concepto particular para medir la relación entre variables aleatorias y que las distribuciones marginales más la estructura de correlaciones no definen las distribuciones conjuntas
- Hemos definido el concepto de cópula que nos podrá ser util para simular variables aleatorias con distribuciones marginales determinadas, además hemos visto algunos ejemplos
- Hemos visto algunas medidas nuevas de la relación entre variables aleatorias, la rho de Spearman y la Tau de Kendall, en particular hemos visto otra forma de estimar la correlación lineal.
- En relación al riesgo de crédito hemos visto el modelo de Merton como un caso particular de cópula, además hemos visto que las correlaciones de default son muy diferentes de las correlaciones con las que habitualmente trabajamos.

BIBLIOGRAFÍA:

1. **Linear Correlation Estimation**, Filip Lindskog - http://www.math.ethz.ch/~lindskog
2. **Modelling Dependence with Copulas and Applications to Risk Management**, Master Thesis by Filip Lindskog
3. **Understanding Relationships Using Copulas**, Edward W. Frees and Emiliano A. Valdez
4. **Correlation And Dependence In Risk Management: Properties and Pitfalls**, Paul Embrechts, Alexander McNeil, and Daniel Straumann
5. **Correlation: Pitfalls and Alternatives**, Paul Embrechts, Alexander McNeil, and Daniel Straumann
6. **Copulas for Finance: A Reading Guide and Some Applications**, Eric Bouyé, Valdo Durrleman, Ashkan Nikeghbali, Gaël Riboulet and Thierry Roncalli
7. **On Default Correlation: A Copula Function Approach**, David X. Li